

AD 626145

NOTS TP 3934
COPY 77

LIBRARY INFORMATION SEARCH AND RETRIEVAL DATA SYSTEM (LISARDS)

by

J. L. Zaharias
Systems Development Department

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION		
Hardcopy	Microfilm	
\$2.00	\$0.50	41.21
ARCHIVE COPY		

ABSTRACT. A series of report literature-searching programs has been written for the IBM 7094 computer at the U. S. Naval Ordnance Test Station, China Lake, for use in the technical library to augment manual literature-searching techniques. In addition to the programs written in direct support of the searching activity, several programs have been written to decrease the time needed for sorting other library records. These auxiliary computer programs include the automatic posting of new reports in the manual coordinate index according to descriptors, the printing of the authority file used by indexers, and the preparation of 3x5-inch catalog cards. This report discusses indexing and literature-searching routines, and the evolution of the computer retrieval system. Programs written for the computer, and the procedures for using them, are presented in Appendix A through F.

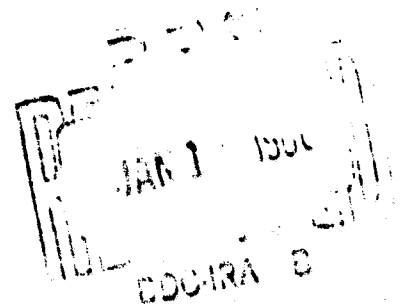
Permission to reproduce
Clearinghouse for Federal Scientific
and Technical Information given by
U. S. Naval Ordnance Test Station,
China Lake.



U. S. NAVAL ORDNANCE TEST STATION

China Lake, California

November 1963



**Best
Available
Copy**

U. S. NAVAL ORDNANCE TEST STATION

AN ACTIVITY OF THE BUREAU OF NAVAL WEAPONS

J. I. HARDY, CAPT., USN
Commander

WM. B. MCLEAN, PH.D.
Technical Director

FOREWORD

This report was written to document the progress made in literature search and retrieval computer programs being prepared at the U. S. Naval Ordnance Test Station for use in technical libraries to augment presently used techniques.

Programs written for the IBM 7094 computer, the procedures for using them, and auxiliary programs for use in conjunction with manual methods are presented.

This work is part of the data computation development being carried on at NOTS in various applications of data automation.

Released by
LEE E. LAKIN, Jr.
Head, Assessments Division
17 September 1965

Under authority of
IVAR E. HIGHBERG, Head
Systems Development Dept.

NOTS Technical Publication 3934

Published by	Systems Development Department
Manuscript	30/MS-738
Collation	Cover, 19 leaves, DD Form 1473, abstract cards
First printing	165 numbered copies

CONTENTS

Introduction	1
Indexing Procedure	1
Literature Searching	2
Background	2
Evolution of the Computer Information Retrieval System	3
Conclusion	6
Appendixes:	
A. Prepare and Edit Master Report Tape	7
B. Correct and Change Master Report Tape	10
C. Updating a Master Descriptor Tape	12
D. Final Search Program	21
E. Print Authority File	29
F. Computer System Control Cards	33

INTRODUCTION

A major problem encountered in handling any large collection of items is the proper indexing, sorting, and retrieval of the items. Many cataloging and cross-referencing systems are used to provide orderly storage and maximum use of the items. However, libraries (and particularly technical libraries) are faced with multiple indexing and retrieval problems since reports and documents can, and usually do, fall into more than one subject category. Therefore, extremely detailed and comprehensive indexing and filing are required in order to insure the efficient retrieval of material in a burgeoning number of subjects, especially in the fields of research and development.

The procedure adopted by the technical library at the U. S. Naval Ordnance Test Station (NOTS), for cataloging the vast number and variety of reports and documents housed in its three facilities, is the inverted file indexing system using the 'uniterm' cards developed by Mortimer Taube to provide the user and the indexer with a large number of indexing terms, and by later converting it to a combination automatic/manual information retrieval system.

INDEXING PROCEDURE

With the uniterm card system, as a report is entered into the collection and given an accession number, it is analyzed and placed into several subject categories according to descriptors that describe the contents of the report. The descriptors are taken from a NOTS-developed authority file containing a comprehensive list of descriptors, synonyms, and cross-reference items. If it becomes necessary to use a descriptor not already entered in the file, the term must first be reviewed and approved by the library staff before it is entered into the control authority file as a functional descriptor.

LITERATURE SEARCHING

Individual cards are prepared for each descriptor, and the accession numbers of all reports falling into that specific subject category are posted on the cards. Thus, a searcher wishing to find all the documents in the library dealing with submarines, for example, would select the coordinate index cards using the descriptor, 'submarines'. If the user wishes to limit the reports to those on submarine detection, he selects the descriptor file cards for both 'submarines' and 'detection'. The accession numbers appearing in common on the cards will be the accession numbers of all reports having to do with submarine detection.

This process can be followed to select reports covering more detailed information on submarines (or any subject) such as design, operation, warfare, testing, and maintenance by using the correct descriptor cards. However, the process becomes cumbersome, inefficient, tiring, and very time-consuming when the search is made manually, using cards having many accession numbers. To alleviate this situation, it was decided to use an electronic computer to aid in literature searching; i.e., information retrieval.

BACKGROUND

At the time the new concept of using an electronic computer was first applied to information retrieval at NOTS, there were about 6,000 descriptors in the coordinate index system. These were put into alphabetical order and each was assigned a descriptor number; then, the descriptor numbers of the entire system and their respective report accession numbers were put on a machine-usable magnetic tape. As new descriptors were added to the system, each was given a number one higher than the preceding descriptor (a practice still in effect today) and the new numbers were added to the end of the tape.

The machine was programmed to read the desired descriptor numbers from punched cards, correlate them with the magnetic tape information, compare the report accession numbers, and print the results.

Although the output from this searching process was an accurate and complete list of accession numbers common to the compared descriptors, it was necessary for the searcher to go to the shelf list to find the complete information for each report accession number on the computer list. In addition to this complicated process, other, seemingly insurmountable, restrictions tended to make it appear that literature searching by computer was impractical. The restrictions included the large number of

tape reels that had to be stored because of the mass of data and the relatively low recording density of the tape; the internal storage capacity of the computer limited the number of reports that could be associated with any one descriptor to 6,750; and only eight descriptors could be used for any individual search. Therefore, the computer retrieval service was discontinued, although the punching of cards was continued so that the technique would not be lost.

However, with the development of new computers, tape drives, techniques, and memory capabilities, it was found that the restrictions were not valid. Multiple input and output reels of magnetic tape could be easily handled, and, with the new equipment came a higher degree of reliability in tape drives, a higher quality of magnetic tape, and a quadruple increase in the data-recording density of the tape. It was found also that the maximum of 6,750 reports per descriptor did not hamper the process since, even today, the maximum number of reports attached to a single descriptor is under 2,000; and the limit of only eight descriptors for any one search is well within the average two usually needed (the maximum number of descriptors used for a search to date has been only five).

EVOLUTION OF THE COMPUTER INFORMATION RETRIEVAL SYSTEM

Since the only valid reason for lack of use of the service apparently was the necessity for the searcher to go back to the shelf list for final report selection, when the IBM 7094 computer was installed at NOTS, it was decided to reactivate and improve the output of the information system and to add to the number of operations that could be handled by the computer. While this in no way would solve all the retrieval problems encountered nor do away with personnel needed in the library, it would speed up and immeasurably improve the retrieval process.

Improvements in Computer Searching

To improve information output, the list of report accession numbers formerly presented to the searcher is now augmented with more extensive and complete information about the reports obtained in the search. The added information includes such items as report title, issuing agency or corporate author, name of author, and security classification of the report.

This improvement was achieved by taking all the desired information from the shelf list and manually keypunching it onto cards to be fed into the computer where the information is transcribed on magnetic tape. This tape, used with the proper computer program, gives the searcher all the information he formerly had to obtain from the shelf list after he had been given a list of report accession numbers. The master magnetic tape file is updated as new reports enter the collection, just as the shelf list is kept up-to-date in the manual system.

As a by-product of the computer program that keeps the machine coordinate index file up-to-date, a new card is made by the computer for each descriptor that has been changed and for each new descriptor added to the manual coordinate index. Library personnel place the new cards in their proper places in the manual file and destroy the old cards for the updated descriptors.

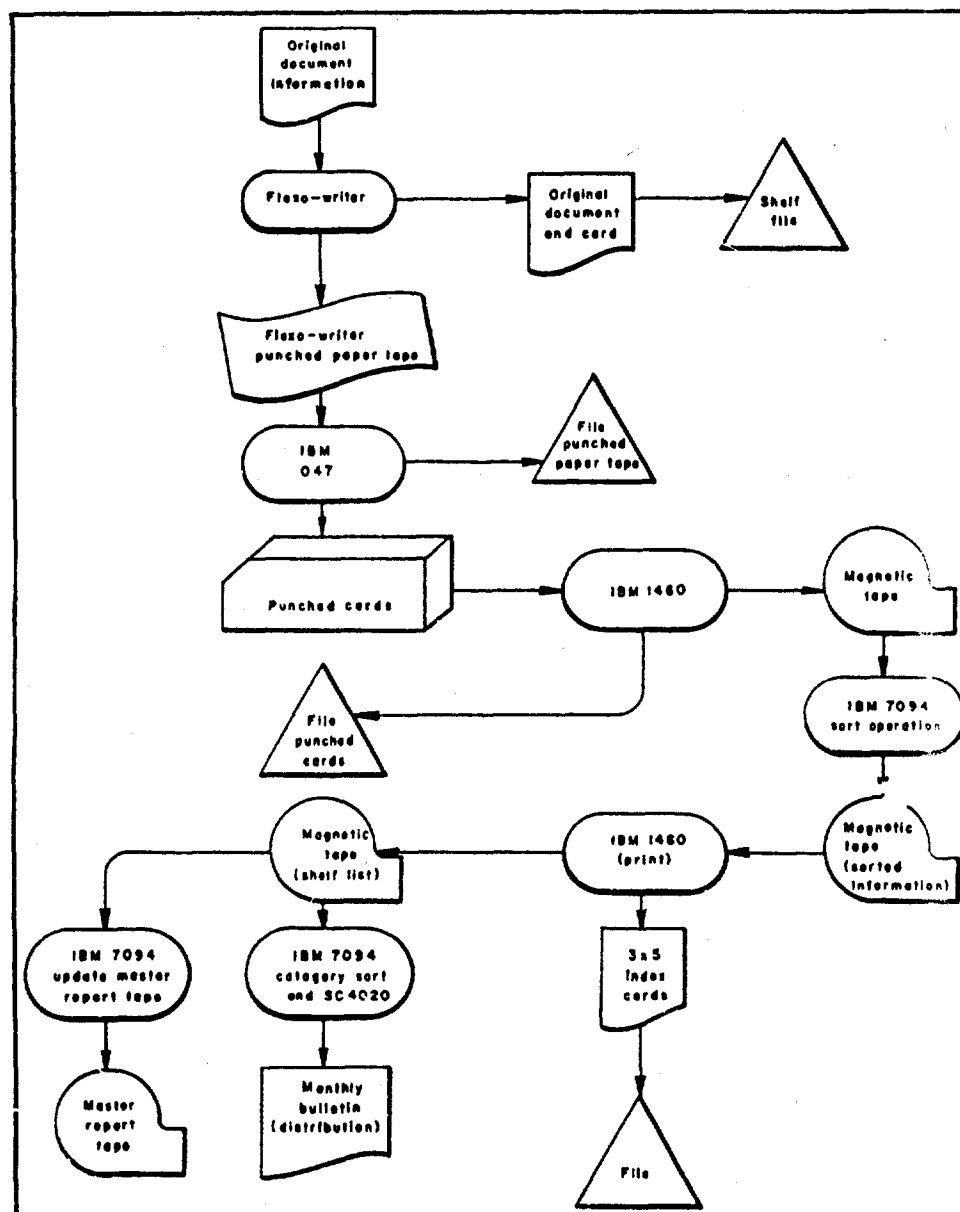
Another time- and labor-saving scheme that has evolved from the retrieval system is the preparation by computer of the 3x5-inch cards used in the card catalog of reports. The China Lake library staff prepares all filing cards for the collection of reports at the main library and also for those at the Pasadena Annex and China Lake Pilot Plant libraries. The cards contain all descriptive information and are filed according to the following groups in the catalog:

1. Shelf list cards, filed in ascending order by accession number.
2. Issuing agency or corporate author cards, filed in alphabetical order.
3. Personnel author cards, filed in alphabetical order (a card is prepared for each author of each report).
4. Government contract cards, filed according to contract number.
5. Series cards, filed according to series number.

To produce the many cards needed for this multiple card cataloging, a Flexowriter is used to produce a punched paper tape at the time of the original descriptive cataloging. The paper tape is fed into an IBM 047 paper-tape-to-card converter, which prepares cards for input to an IBM 1460 computer program. The output of the 1460 at this point is a magnetic tape that contains one physical tape record for each of the five types of cards needed. The record contains all the information that appeared on the original typed card, plus the number of copies needed to supply all three libraries and the identification of the general fil-

ing category to which the report belongs. This information is then sorted by the IBM 7094 program and transcribed on another magnetic tape which is fed back into the IBM 1460 for the final printout of the 3x5-inch cards, which are produced in the proper number and in the proper order for filing.

During the printing process, the shelf list records are simultaneously rewritten on another magnetic tape for use in keeping the master report tape up-to-date and to produce a monthly information bulletin that lists new reports added to the collection. The following diagram shows the flow of information in the automated production of the catalog cards.



CONCLUSION

The evolution of computer information retrieval at NOTS has produced a combination automatic/manual system that provides a more effective service for the users of its three technical libraries. In addition, library personnel have been relieved of many tedious, repetitive, and time-consuming tasks, enabling them to add to the efficiency of the entire library function and to supply the thorough, less-hurried, and more personal attention so often needed by library users.

Appendixes A through E describe various information-retrieval computer programs; Appendix F describes the computer system control cards.

Appendix A

PREPARE AND EDIT MASTER REPORT TAPE
(Programs 00712 and 00723)PREPARE MASTER TAPE
(Program 00712)

This program is used to add report information to the master magnetic tape, which contains information on all reports in the NOTS technical libraries. Program inputs include the master tape and the information to be added to the master tape from punched cards or auxiliary magnetic tapes.

The master tape is written in the binary coded decimal (BCD) mode, the first physical record on which is an identifying symbol which contains the word, 'MRT 01'. The remaining physical records are 4,000 computer words long. The records containing the descriptive information for each report, are stored numerically by accession number. After each piece of report information is recorded, a logical end of record word (L.E.R.) is written and, at the end of the physical record, a physical end of record word (P.E.R.) is written. The last record of the entire file contains a physical end of file word (P.E.F.) after all the new information has been written. After the last record on the magnetic tape, an end of file is written. The master report tape is referenced by the computer program by its designated logical name, 'SYSBR6'.

Input Cards

The program input cards contain the following information and are punched according to the following format:

<u>Column</u>	<u>Description</u>
1	Sequence number of card
2-6	Report number
7	'P' if report is a progress report; blank if it is not
8	'L' if report is a NOTS report; blank if it is not
9	Security classification of report 'C' for Confidential, 'S' for Secret, blank if Unclassified

With the exception of column one, the above columns are punched with the same information for all the cards for one report accession number.

Columns 10 through 72 contain descriptive information about the report in the following order:

Issuing agency (corporate author)

Title

Miscellaneous (such as volume number,
number of pages, publication date,
etc.)

Name of author(s)

An asterisk is placed between the different groups of information.

Input Deck

The cards in the system control card deck are described in Appendix F of this report. They are placed in the following sequence:

1. Job submittal form
2. System job card
3. System load card
4. A blank card
5. Program deck marked 00712
6. Data control card (the NOEDIT feature is used on this card)
7. Data cards follow in numerical order

Output From the Program

Since the program adds the new reports to the end of the input tape, the program output tape is the complete listing of reports in the three libraries.

EDIT PROGRAM (00723)

An edit program can be used in conjunction with the master tape program to check for accession numbers out of order or missing from the master tape. To use the edit program, the input deck sequence is changed to:

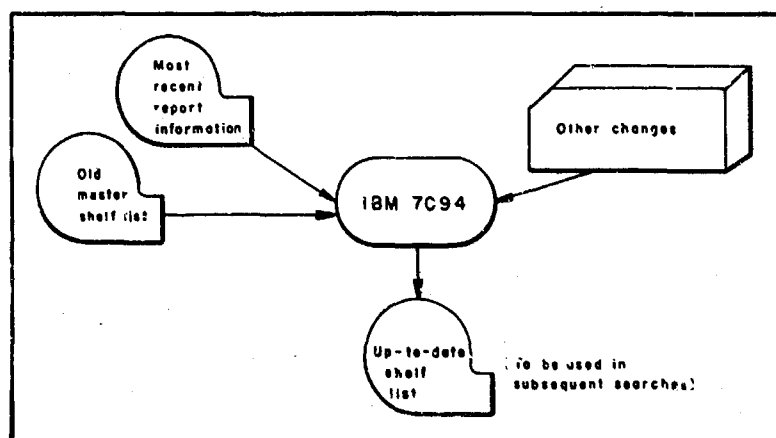
1. Job submittal form
2. System job card
3. System link card (new insertion)
4. System load card
5. A blank card
6. The edit program deck numbered 00723 (the data control card-- or NOEDIT--feature is dropped)

Output From the Edit Program

The edit program tells the user what reports are on the master report tape and shows any errors that might be on the tape. This program checks accession numbers to make sure that each is in numerical order and that no accession number is missing. When either type of error exists, the program prints the corrected or added accession number, together with the information from the original report card, immediately in front of the next higher accession number on the tape. The output is then used to make the necessary changes and corrections to the master tape.

NONSYSTEM MAGNETIC TAPES

A nonsystem magnetic tape is also used with the master tape as an input/output tape referenced by the program by its logical name, 'SYSBR6'. The nonsystem tape is unique to an individual program and is used to store reusable information. On the job submittal form, BR6 is written in the proper place, with the correct reel number of the current master report tape, and circled to insure that the tape will be saved for future use, see flow chart of changing shelf list tape.



Appendix B

CORRECT AND CHANGE MASTER REPORT TAPE
(Program No. 00724)

Program 00724 was written to change and update the report information on the master report tape described in Appendix A. The changes may be necessary because of recataloging, errors in the initial information supplied, or a change in security classification. The inputs to the program are the master report tape and the punched cards which indicate the changes.

The logical name given to the master input tape for this program is 'SYSAR4'. It is read by the program and corrections indicated by the input cards are made. A new, updated master report tape is written in the process of changing the information. The new master tape then is known by its logical name 'SYSBR6'.

Input Cards

The three types of input cards to this program (deletion, insertion, and information cards) are punched in the formats given in the following sections.

Delete Control Cards. A delete control card, used to delete an accession number from the master tape, is prepared by punching in the following information:

<u>Column</u>	<u>Description</u>
7-12	The word 'DELETE' is punched in these columns
19-24	The accession number of the report to be deleted is punched here. The leading zeros must also be punched.

The delete card is used to delete words that may appear as report information by mistake. All information is deleted beginning from the group of characters punched in columns 19 to 24 up to, and including, the logical end of record word (L.E.R.) on the magnetic tape.

Insert Control Cards. The insert card is used to signal that all the succeeding cards shall be inserted on the magnetic tape. It is punched as follows:

<u>Column</u>	<u>Description</u>
7-12	'INSERT' is punched here

All the following cards up to a 'DELETE' card or to the end of data will be inserted in their proper places on the master tape.

Information Input Card Format. This type of card is used to insert information and is punched using the format described for input cards in Appendix A.

Input Deck Preparation

System control cards are described in Appendix F. The input sequence of the control cards for this program are the same as for Program #00712 described in Appendix A, using the data control card--or NOEDIT option--in position 6, followed by the data cards in numerical order. In addition, the last two cards must be punched 999999 in columns 1 through 6.

Program Output

The output from Program #00724 is a corrected and updated master report tape having the same format as the input to the program. The program calls the output tape by its logical name, 'SYSBR6'.

Additional Program Usage

The edit master report tape program, No. 00723, can also be used with this program by making the card sequence changes described in Appendix A.

Nonsystem Magnetic Tapes

Two nonsystem magnetic tapes are used by this program. The input tape is the master report tape, called by the computer program by its logical name, 'SYSAR4'; the output tape is the updated master report tape called by the program by its logical name, 'SYSBR6'.

Appendix C

UPDATING A MASTER DESCRIPTOR TAPE (A support computer program in an information retrieval system)

INTRODUCTION

This computer program was written as part of an entire search system for the Technical Library. Its prime function is to keep current the master descriptor tape which contains the entire coordinate index information in computer-usable form. Its secondary function is to prepare coordinate index cards for the manual files kept in the Technical Library. The coordinate index contains all the descriptors or uniterms, each of which has been assigned a number in the system; the title or name of the descriptors; and all of the accession numbers of the documents under each of the descriptors.

PURPOSE

This program was designed to perform the same operations as the posting clerk in the manual system of updating the coordinate index. These operations are

- Add new descriptors to the system or
change existing titles in the system

- Add new descriptors to the system along
with their respective accession numbers

- Add new accession numbers to descriptors
which are in the system

- Delete accession numbers from descriptors
in the system

- Delete descriptors and all relative information about the descriptors from the
system

OPERATION

Using the above functions and the old master descriptor tape, the computer program will prepare a new master descriptor tape to be used for searching the literature in the library using another computer program. At the time of preparing a new master descriptor tape, another magnetic tape is prepared to be used to produce the coordinate index cards for the descriptors which have been changed or added to the master

descriptor tape during this updating process. This magnetic tape is processed by a second computer program which prepares the information for printing in the same format as the coordinate index cards. The entire task of updating the master descriptor tape and producing new cards for the coordinate index file is actually two computer programs run one after the other. The first program updates the master descriptor tape and the other produces the coordinate index cards.

Inputs

The inputs for the first computer program are

The master descriptor tape

The punched cards which indicate the changes or additions to the descriptor titles

The punched cards which indicate the changes to be made to the file of descriptors already on the master descriptor tape

The master descriptor tape contains all the coordinate index information in computer-usable form. Each descriptor is on the magnetic tape as a record of information. This record contains the descriptor number, the title of the descriptor, and all the accession numbers of the reports which have been assigned to this descriptor. Along with each accession number there is an adjacent computer word showing the security classification of the report and if the report is a progress report. The descriptors are on the magnetic tape in numerical order according to descriptor number. The accession numbers under each descriptor are also in numerical order.

The punched cards which contain the changes to the titles of the descriptors are punched in the following format.

<u>Column</u>	<u>Description</u>
1	This column contains a six (6) which is used as a control punch
2-7	This data field contains the descriptor number with leading zeros punched
8	This data field contains an asterisk (*) if the document from which this card was punched contains a scope note. It is left blank if no scope note is present.

- 9-68 This data field contains the title of the descriptor. The entire descriptor title may be up to 180 characters in length, including spaces and special characters. If this is the case, up to three cards may be used for one descriptor title.
- 72 This data field contains the sequence number of the punched card. It may be a 1, 2, or 3.

The punched cards which contain the changes to be made to the file of descriptors on the master descriptor tape all have the same format except that the control punch in column one of the card defines what is to be done with the information on each card. The control punches and card format are

<u>Column</u>	<u>Description</u>
1	Control punch (defined below)
4-9	This data field contains the descriptor number with leading zeros punched
13	This data field contains an "S" if the report whose accession number appears in this card has a security classification of Secret. It is left blank if any other classification is attached to the document
14	This data field contains a "P" if the report whose accession number appears on this card is a progress report, otherwise it is left blank.
17-22	This data field contains the accession number of the report with leading zeros punched
25-72	This data field contains the descriptor title

The four control punches used by the computer program are

<u>Control</u>	<u>Description</u>
1	This control punch in column one indicates to the computer program that the accession number punched in the card is to be added to the already existing descriptor record whose number appears on the card, along with all other information punched in the card
2	This control punch in column one indicates to the computer program that the accession number in the punched card is to be deleted from the descriptor record whose number appears in the punched card.

- 3 This control punch in column one indicates to the computer program that the entire descriptor record whose number appears on the card is to be deleted from the master descriptor file. All information pertinent to the descriptor will be deleted from the magnetic tape by this card.
- 4 This control punch in column one indicates to the computer program the descriptor whose number is punched in the card will be added to the magnetic tape file along with the accession number on the card. If there is more than one accession number to be filed under a descriptor number, there will be one card for each accession number with this same control punch in column one.

Output From the Updating Program

One of the outputs which must be mentioned here for continuity is the magnetic tape which contains all the changes that were made to the master descriptor tape. This is just one of the outputs from the program. It will be used by the second program which processes the tape. Other outputs are described on page 18 of this Appendix.

Input Cards to the Second Program

The input cards to the second program which prepares the coordinate index cards are punched in the following format.

<u>Column</u>	<u>Description</u>
1	A one (1) is punched in this column. It is used as a control punch to the program.
2-7	The number of the descriptor which is to mark the beginning of the printing is punched in this data field with leading zeros.
8-13	The number of descriptors which are to be processed beginning with the descriptor punched in the previous data field. The number is punched with leading zeros and is right justified.
14	This column contains another control punch to the program. If it contains a zero or a blank, the instructions indicated by the other data fields in this card will be followed. The program will search the input tape to find the descriptor punched in columns 2-7 of this card and process as many descriptors following as indicated in columns 8-13.

If this column contains any punch other than a zero or a blank, all the descriptors which are on the magnetic tape beyond the point at which the tape is now positioned will be processed.

As many cards as are needed to perform the complete operation may be used with this program. This data input configuration allows the user of the computer program to print selected sections of the input tape or the entire input tape.

Input Card Order

The order for the input cards in the updating of the master descriptor tape and the production of the coordinate index cards is shown below.

- a. Job Submittal Form (11ND-NOTS-10462/19)(Rev. 3/65)
- b. System job card
- c. System link card
- d. System load card
- e. A blank card
- f. The first program deck is placed in this position
- g. System data control card. The "EDIT" option is used with this computer program.
- h. A blank card
- i. Format cards #1 through 6 are put in here. These cards, which are described on page 18 of this report, are used to define the format of the input cards for this program.
- j. The input cards described on page 13 which indicate the changes to be made to the titles of the descriptors are placed here.
- k. A card with a "12" punch in column one is placed here to indicate the end of the change cards.
- l. The input cards described on page 14 which indicate the changes to be made to the master descriptor tape information are placed here. These cards are

placed in numerically ascending order according to descriptor number.

- m. A card with the following format is placed here.

<u>Column</u>	<u>Description</u>
1	A five (5) is punched in this column for control purposes.
4-9	777777 is punched in this field for control purposes.
	n. A card with a "12" punch in column one is placed here to indicate the end of this group of data input, followed by a card with a comma (,) in column one to indicate the end of all data for this computer program.
	o. System link card. This card marks the beginning of the second computer program input deck. This second computer program produces the coordinate index cards.
	p. System load card.
	q. A blank card
	r. The second computer program deck is placed here.
	s. System data control card. The "EDIT" option is used here.
	t. A blank card.
	u. The #7 card with a format number of slash (/) is placed here. It is defined on page 18 of this report.
	v. The input cards to the second program are placed here. They are described on page 15.
	w. A card with a comma (,) in column one is placed in this position to indicate the end of all data for both programs.

Description of the Format Cards

The format cards for these two computer programs are punched as follows according to the specifications given in the Variable Field NOTS Input Editor by Helen Morrin, dated 1 November 1959. All format cards begin in column one.

1. /4/X,1,1/H,4,6/H,17,6/H,13,2
2. S4/X,1,1/H,4,6/H,17,6/H,13,2/
3. T4/X,1,1/H,4,6/H,17,6/H,13,2/
4. U4/X,1,1/H,4,6/H,17,6/H,13,2/
5. V2/X,1,1/H,4,6/
6. W11/H,2,6/10H,9,61/
7. /3/H,2,6/X,8,6/X,14,1/

NONSYSTEM MAGNETIC TAPES

There are four nonsystem magnetic tapes used in the execution of this program. One is used as the input tape; two as output tapes; and another as an intermediate tape which is not saved for future use.

Input Magnetic Tape

The input tape to this computer program is the master descriptor tape which is to be updated with the new material. Its symbolic name is SYSBR4. This magnetic tape is saved in the reserve pool of tapes for a period of time after the updating has taken place.

Intermediate Magnetic Tape

An intermediate tape which is made at the same time as the updating of the master descriptor tape is used in the second program to produce the coordinate index cards. The symbolic name of the magnetic tape is SYSBR7. This tape need not be saved.

Output Magnetic Tapes

The produced output tape from the first computer program is the updated master descriptor tape. Its symbolic name is SYSAR4 and is always saved for future use. A copy is made as soon as possible for use in case of emergency.

The output tape from the second computer program is the tape which contains the coordinate index information. The symbolic name of the magnetic tape is SYSAR1. The magnetic tape is printed on the IBM 1401 using the special forms for the technical library and the special IBM 1403 carriage tape. Two copies of this listing are produced for the coordinate index files, one for the China Lake technical library and the other for the Pasadena library. After printing, the magnetic tape is returned to the pool of magnetic tapes for reuse.

OUTPUT MESSAGES

The computer program is written to produce messages so that the user of the program will have some indication of what has happened during the updating process and be able to make the necessary notations and corrections.

Messages

1. DESCRIPTOR NUMBER XXXXXX DELETED. This message is self-explanatory. An indication of this event should be noted in the file of descriptors for future reference.

2. NO REPORT NUMBER YYYYYY IN DESCRIPTOR NUMBER XXXXXX. This message indicates that the report number shown does not exist under the descriptor number shown. This happens in the event that a report number is being deleted from a descriptor and the report number was not under that descriptor on the magnetic tape. This could also indicate an error in transcribing the data.

3. NOT AN INSERTION FOR DESCRIPTOR XXXXXX YYYYYY. This message indicates that an attempt is being made to insert a report number to a descriptor which is not on the magnetic tape. This may indicate that the input card is out of order or that the card has been punched incorrectly. This message needs close scrutiny before the next updating process.

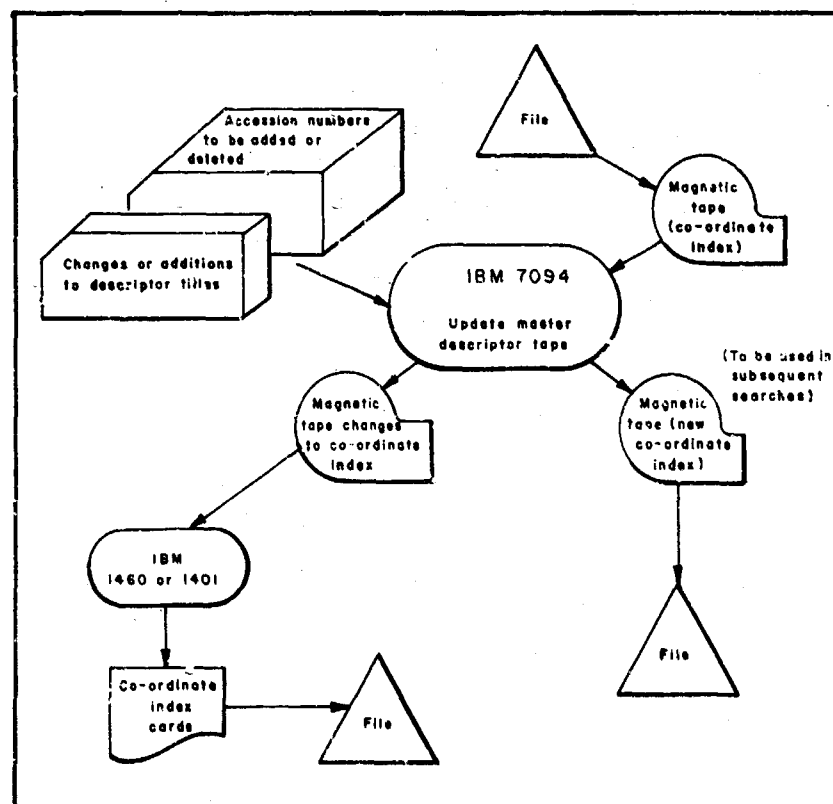
4. SEEMS LIKE WE RAN OUT OF REPORTS ON DESCRIPTOR XXXXXX. This message indicates that all report numbers were deleted from the given descriptor record. The computer program has deleted the entire descriptor record from the magnetic tape. This should be noted in the manual descriptor file for future reference.

5. THIS 4 IN COL. 1 SHOULD BE A 1 ON DESCRIPTOR XXXXXX. This message indicates that the user is attempting to insert a descriptor record which is already in the magnetic tape file. This message needs close scrutiny and remedial action. Further investigation will probably show an error in transcribing the descriptor number from one manuscript to another.

6. INVESTIGATE THE FOLLOWING DESCRIPTOR INFORMATION XXXXXX (descriptor title). This message usually indicates that the descriptor indicated was not on the magnetic tape or that the input card was out of order. The input card(s) should be removed from the input deck for insertion in a future updating computer run.

7. LARGEST NUMBER OF REPORTS NNNNNN APPEARS ON DESCRIPTOR NUMBER XXXXXX. This message informs the user what descriptor record is the largest on the entire magnetic tape. This information is used in preparing the correct number of locations in memory for the computer program. No action need be taken except by the programmer--if the limit of 2000 is exceeded.

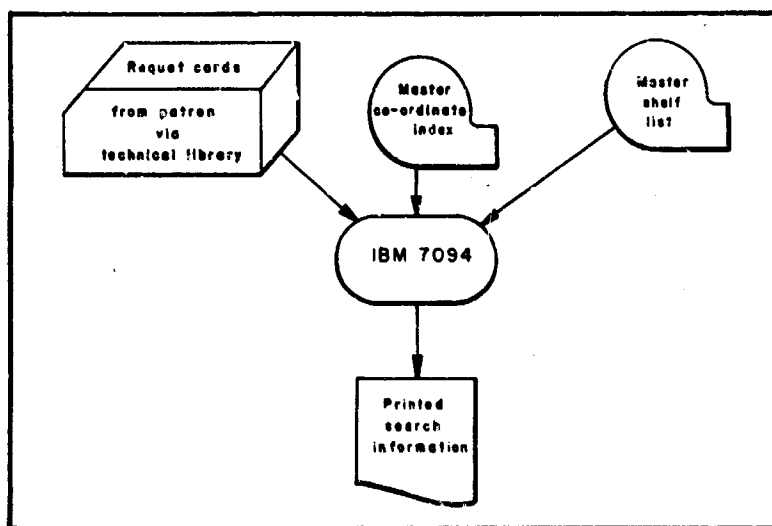
Following is an overall flow diagram of the entire system, including program.



Appendix D
FINAL SEARCH PROGRAM
(Program No. 00722)

PURPOSE

The main object of this program is the production of a bibliography of reports on a specific subject. In this process, which is commonly called coordination, the reports under two or more subjects or terms are compared to find those reports which are common to the terms or descriptors under consideration. The searching process flow is diagrammed below.



With the use of controls, which are read by the program from the input cards, the requester can obtain a complete bibliography of all reports under a given term, or a bibliography of all reports under a given term starting with any report number under the term. During the coordination process, the point may be reached where there are no common reports to be found, that is to say that the request has become too specific for the system. In such cases, the requester may exercise the option to have a list printed of all the reports found up to this point by using the control fields in the input cards.

PROCEDURE

The patron who wishes to obtain information from the literature on file in the library consults a reference librarian and descriptors or uniterms covering the subject are selected for use in the search. Forms, which contain information needed by the program to produce the desired bibliographies, are prepared by the reference librarian. The information from the forms is transcribed to the input cards using a keypunch.

INPUT CARDS

The three cards produced by this keypunching operation contain information identifying the particular focal point in the search. This identifying information is punched into the first input card and is printed as the title of its bibliography. The second input card contains the number for each of the descriptors involved, along with the control punches for any or all of the options in the computer program. This complete information to the program is punched into two input cards, which comprise input for one question, along with a control card, which contains a 12-punch in column one. The control card is placed after the second card of input to indicate the end of a question. As many groups of cards as needed can be used in one pass of the computer. One group of cards, including the last control card, represents a question.

By using the control fields in the second input card of a question, the patron may exercise options which are in the program. The columns in the input cards that are used as signals to the program are described below along with the format of the input cards.

ORDER OF INPUT CARDS

The input to the computer is punched cards and is submitted to be run in the following manner.

A. The first set of cards is the job submittal form which is filled in with all the required information for the computer operators. This is a cover card, unique to the installation, and is used to keep track of the job as it is processed.

B. The second card is a punched IHM card called the "job card". It contains, among other things, accounting information, which allows the charging of time on the computer.

C. The next group of cards is called the "program deck". This deck contains all the programmed instructions telling the computer how to read and process the information given to it and how to print the output information. The program deck also contains control cards that govern the format of the input cards.

D. The next cards are the input cards to the program. They are placed in groups of three cards, each group representing a question. A description of the cards is given here.

a. Card Number 1, The Identification Card

<u>Columns</u>	<u>Punches</u>	<u>Description of the Field</u>
1	8	Control punch
2-67	Alpha-numeric	The title or identifying information written on the form prepared by the reference librarian is punched here. This will be printed as a heading or title to the bibliography produced by the question.
68-70	Three digits	These columns contain the request number, which is assigned by the librarian.
71-72	Two digits	The number of data fields designated as D_n in the next card, which will be used in the question, is punched here.

b. Card Number 2, Descriptor Number and Control Card

<u>Columns</u>	<u>Punches</u>	<u>Description of the Field</u>
1	7	Control punch for the input of data
2-7	Six digits	D_1 - The first descriptor of the question
8	One digit	F_1 - The first control punch for program
9-14	Six digits	D_2 - The second descriptor of the question
15	One digit	F_2 - Second control punch for the program
16-21	Six digits	D_3 - The third descriptor of the question
22	One digit	F_3 - Third control punch for the program
etc.	etc.	etc.

This format is used until all descriptors needed for the question have been punched on the card. As many as ten descriptors and ten control punches can be placed in one card. If there is a need for more than ten descriptors on any one question, a second descriptor card, having the same format as the descriptor card described above, can be used.

c. Card Number 3, End of Group Card

This card contains a 12-punch in column one and is used to indicate to the program that this is the end of a group or question

By using control punches in the second input card of the question (the descriptor number and control card), the patron can exercise program options. The columns (fields) in the input card are described below:

If the control field, F_n , is a blank or zero, the program will attempt to coordinate the descriptors indicated by the input cards and, if successful, will produce a bibliography. If unsuccessful, the program will print a message giving the reason for the failure

If the control field, F_n , is a one, the data field following (D_{n+1}) will be treated as a report number. The ensuing bibliography will contain all the reports on the descriptor or descriptors that have a report number greater than the report number punched in the data field D_{n+1} . This option can be exercised when obtaining reports from a single descriptor or from a coordination of several descriptors.

If the control field, F_n , is a two and no common reports are found after the reports under the descriptor in data field D_{n+1} have been examined and compared with the reports under the descriptor in data field D_n (in the case of just two descriptors), a bibliography of the reports under descriptor D_n will be printed. If this condition occurs on the n th descriptor while doing a coordination of several descriptors, the ensuing bibliography will contain those reports obtained while coordinating the previous descriptors.

E. A card with a comma in column one is placed at the end of all the data to indicate the end of the data cards for this entire search.

The program uses five magnetic tapes other than the computer system tapes. The program refers to these magnetic tapes by logical names. On

the job submittal form, the submitter of the program writes the logical names of the magnetic tapes in their proper place. Both the master descriptor tape and the master report tape are referred to by logical names and by reel numbers. The logical name and reel number are circled and placed under the section called "Input Tapes". The circling indicates that the tape is to be saved. The reel number may be obtained from Code 3037 since the numbers change as new magnetic tapes are made while updating.

The logical names of the magnetic tapes used in the program are as follows:

- | | |
|-----|---|
| BR4 | The master descriptor tape. Its logical name and reel number are written on the job submittal form in the area marked "Input Tapes". Both the logical name and reel number are circled to indicate to the operators that this magnetic tape is to be saved. |
| AR4 | The master report tape. The same rules are applicable here as with the master descriptor tape. |
| AR7 | These logical tape names are written on |
| BR5 | the job submittal form under the area |
| BR6 | marked "Utility Tapes". |

This program can be broken into several phases or sections for ease of understanding. In fact, the program was initially written in phases to facilitate the checkout of the problem.

Phase One

1. The input cards, which were prepared by the reference librarian, are read into memory. These input cards are written on a magnetic tape for future use. Each group of input cards, or request, is written as a unit on the magnetic tape. This unit is commonly called a record.

2. The descriptor numbers read from these input cards are sorted in numerical order into a table in the memory of the computer.

3. After all the input cards are read and recorded on magnetic tape, and after all the descriptors from these input cards have been sorted, a file mark is written on the magnetic tape to indicate the end of the input card information. This magnetic tape is then rewound and will be used later in the program.

4. At the end of Phase One, a list of all the descriptor numbers needed for the entire computer pass is in memory, and the input cards, as they were read into memory, are on magnetic tape.

Phase Two

1. The master descriptor tape, containing all the descriptors in the system with their respective report accession numbers, is read into memory and the descriptors needed during this computer run are written on another magnetic tape. The needed descriptors, with their respective report numbers, are found on the master descriptor tape from the table of descriptors in memory produced in Phase One. Only one reading is made of the master descriptor tape. If any descriptor in the memory table is not found on the master descriptor tape, the program will print a message to warn the user that this condition has occurred.

2. Each needed descriptor with its related report numbers is written on the magnetic tape as a record. After all the required descriptors have been read from the master descriptor tape and written on the magnetic tape, a file mark is written on the magnetic tape to indicate the end of the information from the master descriptor tape. The newly written magnetic tape is rewound for future use by the program.

3. In addition to the information obtained in Phase One, there is now a magnetic tape that contains the necessary descriptors and report accession numbers.

Phase Three

1. The magnetic tape written in Phase One containing the input card data is read into memory, one question at a time. From this input information, it is known what descriptors are to be used in the question and whether they are to be used to obtain a coordination or simply to obtain a printing of all the reports under one descriptor.

2. The descriptors required to perform the operations requested, along with the report accession numbers attached to the descriptors, are read into memory from the magnetic tape produced in Phase Two.

3. The operations requested are performed by the computer. If a printing of all reports under a single descriptor is requested, the report accession numbers under the descriptor are sorted into a table in memory. The report numbers, along with the question identification from the input cards, are written on a magnetic tape as a single record. If a coordination of two or more descriptors is required, the operation is completed and the resultant report numbers and the question identifica-

tion from the input cards are written on a magnetic tape as a single record. These report numbers are also sorted into a memory table in numerical order.

4. This procedure is followed until all the questions have been processed. Then a file mark is placed on the magnetic tape to indicate the end of the information. The magnetic tape is rewound for future use by the program.

5. At the conclusion of Phase Three, the magnetic tape produced in Phase Two, containing the descriptors and report numbers, can be discarded.

6. At the end of this phase of the program, the following information has been obtained either in memory or on magnetic tape.

- a. A magnetic tape containing the input cards in the same order that they were read.
- b. A table in memory of all the descriptors needed for the completion of the entire computer run.
- c. A table in memory of all the report accession numbers needed for the completion of the entire computer run.
- d. A magnetic tape containing the identification obtained from the input cards along with the report accession numbers needed for each question.

Phase Four

1. The master report tape, containing bibliographic information of all reports in the system, is read into memory. Using the list of report accession numbers in memory produced in Phase Three of the program, each report number needed for the entire search, plus all related report information, is written on a magnetic tape as a record.

2. A file mark is written on the magnetic tape to indicate the end of the report information. This magnetic tape is rewound for future use by the computer.

3. In addition to the information at hand after Phase Three, there is now a magnetic tape containing all the necessary report accession numbers with their respective bibliographic information.

Phase Five

1. The identification obtained from the input cards and the report numbers needed for each question are read into memory from the magnetic tape produced in Phase Three.

2. The magnetic tape produced in Phase Four containing all the report accession numbers and affiliated information is searched to find all the reports needed to answer the first question.

3. The identification for the question, read from the magnetic tape, is printed as the title for the bibliography.

4. The report accession numbers, along with their related information, are printed to produce the desired bibliography.

5. This is done until all questions have been answered.

The program has now completed the task of searching the literature on file in the technical library and producing bibliographies to answer the requests made.

The master descriptor tape and the master report tape are removed from the computer and placed in the tape file for future use. The bibliographies obtained are printed and returned to the requester.

Appendix E
PRINT AUTHORITY FILE
(IBM 1401 Program No. 00730)

PURPOSE

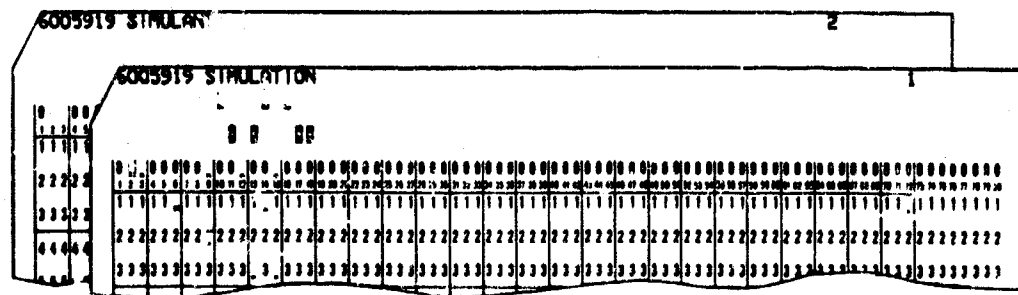
This program was written to provide working copies of the authority file of descriptor titles and cross-references for indexing reports. The manual file is kept on 3x5-inch cards that not only contain descriptor titles and their numbers, but also define and explain the scope of the subject matter. The cards are called 'scope notes'. The information from this file is punched into IBM cards which are filed in alphabetical order. Cards are printed about once a quarter for librarians at China Lake and Pasadena.

INPUT CARDS

The input cards to the IBM 1401 program are punched in the following format.

<u>Column</u>	<u>Description</u>
1	6-control punch
2-7	The descriptor number is punched here with leading zeros. This field is left blank if this reference is to be a cross-reference or a "see also" card.
8	An asterisk (*) is punched here if this reference has a scope note on the original 3x5-inch card. The column is left blank if there is no scope note.
9-68	These columns contain the title of the descriptor, the "see", "see also", or cross-references.
72	This column contains the sequence number of the card. There may be up to three cards for one reference.

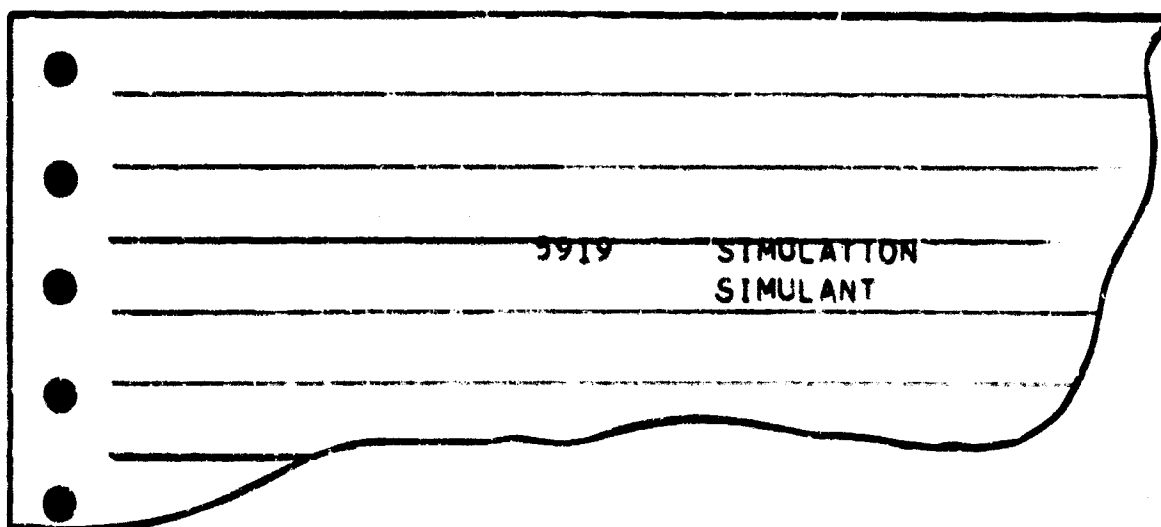
The format of the cards is most important to the final output of the program and since the same format is used in printing the descriptor titles, care should be taken in the presentation of the information to make an easy reading output. An example of the input cards and how this information is finally printed in the authority file list and on the coordinate index cards is given on the next page.



Sample input cards (used to transfer information on magnetic tape).

SIMULATION										5919
SIMULANT										
0	1	2	3	4	5	6	7	8	9	CARD 1
2120	P 2101	P 162	S 3063	3984	2335	316	2407	878	P 2099	
2730	3591	2852	5563	6484	5535	P 3156	3097	2208	3379	
3080	5541	3852	8353	7154	5555	6496	3457	S 2558	5539	
3520	7221	5362	11003	8034	5825	8636	3577	3378	5899	
S 5610	8221	5532	11393	8354	S 6035	10936	P 4357	4418	6109	
	SP 9071	8352	11773	8386	7665	SP14236	10357	5028		

Magnetic tape is then used to print 8x5-inch coordinate index cards.



At the same time, a working copy of the authority file is printed.

OUTPUT

The printed output is spaced so that each reference is separated from references both above and below it. This allows the reader to find his entry with more speed. During printing, each card of information for one reference is printed under the previous line of information. The descriptor number appears only once for each set of cards. An asterisk printed by the first line of information denotes that a scope note is present.

Several copies of the authority file can be printed by using multiply paper on the IBM 1403 Printer.

TIMING

The entire time taken in producing the authority file listing is about an hour and 15 minutes. This includes changing paper and setup time, but does not include filing time.

Operating Instructions

Ready the 1403 printer, spacing down from the top of the page so that the first line will print 5 spaces from the top.

Ready the 1402 reader-punch.

Load the program followed by the input cards.

Press start button to release last cards.

Program Stops

191 indicates the end of the program

189 indicates an error has been encountered in reading the input cards. The last card in the hopper is the one that has been misread. To correct the error, non-process the cards in the card input hopper. Replace these cards with the last card in the output hopper and press the start button.

SPECIAL NOTE

When a card is found that does not have a 1, 2, or 3 in column 72, the entire card image will be printed with three asterisks to the right of the printed line. A card will also be punched so that the operator can make the correction before the next printing is made.

COMPUTER CONFIGURATION

No special features are needed to produce this listing.

Appendix F COMPUTER SYSTEM CONTROL CARDS

JOB SUBMITTAL FORM

Form No. 11ND-NOTS-10462/19 is prepared in triplicate by the user of the program. No carbon is required. A slightly reduced sample form is shown below.

JOB SUBMITTAL FORM - CARD NO. 1 11ND-NOTS-10462 19 (REV. 3-65)										Fill out completely and place all copies in front of program deck.		TIME	SERIAL NO.
SUBMITTED BY			TEL NO.		PROGRAM TITLE			PROGRAM NO.		M.J.D.	AUX.	EST. RUNNING TIME (0:11P)	
CATEGORY (CIRCLE)	TYPE OF JOB (CHECK ONE)					EXPECTED OUTPUT (CHECK)				SECURITY CLASS	TOTAL NON-SYS TAPES USED	LOG IN TIME	
UTL SER 3 DEBUG 5 RUN 6 SYSTEMS 8	SOX	FOR TRAN II	TRACES FOR TRAN	OTHER TAPES	100% ONLY	PROG	PRG	PRINT	PUNCH	TAPES	PLOTS	DATA RESULTS	
	<input checked="" type="checkbox"/>							<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		U U	
												S	
TAPE REQUIREMENTS (System tapes need not be specified) GIVE CHANNEL AND UNIT (Symbolic or absolute)												LOG TIME OUT	
INPUT TAPES Give reel no. also (Circle tapes to be saved)				OUTPUT TAPES (Circle tapes to be saved)				UTILITY (Scratch) TAPES (Utility tapes will not be saved)				CODE NO.	
BR4 Reel #				AR4 Reel #				AR7, BR5, BR6				ES NO.	
REMARKS (Special operating instructions, paper, carriage, tape requirements, etc.)													
Master De- scriptor Tape				Master Shelf Tape				<input type="checkbox"/> SEE OPERATOR REMARKS OTHER SIDE					

SYSTEM JOB CARD

Column	Description
1	7, 8, and 9 are punched in this column for control purposes.
8-10	"JOB" is punched here.
12-16	007XX. This is the program number.
17	, - Field separation.
18-23	Six digit job order number.
24	, - Field separation.
25-27	Three digit auxiliary job order number.
28	, - Field separation

29-31	Estimated running time in hundredths of an hour (no decimal point is punched).
32	, - Field separation.
33-62	Any alpha-numeric information may be punched here, usually consisting of the name of the submitter and his phone number. This identifies the job during its running.
63-66	Submitter's code number with the department in columns 63 and 64; the division in column 65; and the branch in column 66.
67	0 to indicate a scientific job.
68	9 to indicate a production job.
70	The number of nonsystem tapes used by the program is punched here.
72	2 to indicate the job will be executed in Phase 2 of the system.

SYSTEM LOAD CARD

<u>Column</u>	<u>Description</u>
1	7, 8, and 9 are punched here for control purposes.
8-11	"LOAD" is punched here.

SYSTEM LINK CARD

<u>Column</u>	<u>Description</u>
1	7, 8, and 9 are punched here for control purposes.
8-11	"LINK" is punched here.
16-72	Any alpha-numeric information is punched here. The information is used to identify the program and should contain the submitter's name, phone number, and the title of the link being run.

SYSTEM DATA CONTROL CARD

<u>Column</u>	<u>Description</u>
1	7, 8, and 9 are punched here for control purposes.
8-11	"DATA" is punched here.
16-21	"NOEDIT" is punched here if the NOTS input editor is not used during this execution.
16-19	"EDIT" is punched here if the NOTS input editor is used during this execution.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) U. S. Naval Ordnance Test Station China Lake, California 93557		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP --	
3. REPORT TITLE LIBRARY INFORMATION SEARCH AND RETRIEVAL DATA SYSTEM (LISARDS)			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Computer program for literature search and retrieval			
5. AUTHOR(S) (Last name, first name, initial) Zaharias, J. L.			
6. REPORT DATE November 1965		7a. TOTAL NO. OF PAGES 36	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. Station overhead funds		9a. ORIGINATOR'S REPORT NUMBER(S) NOTS TP 3934	
b. PROJECT NO.			
c.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.			
10. AVAILABILITY/LIMITATION NOTICES Qualified requesters may obtain copies of this report from DDC. (Released to Clearing House for Federal and Technical Information.)			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Bureau of Naval Weapons (Department of Navy)	
13. ABSTRACT ABSTRACT. A series of report literature-searching programs has been written for the IBM 7094 computer at the U. S. Naval Ordnance Test Station, China Lake, for use in the technical library to augment manual literature-searching techniques. In addition to the programs written in direct support of the searching activity, several programs have been written to decrease the time needed for sorting other library records. These auxiliary computer programs include the automatic posting of new reports in the manual coordinate index according to descriptors, the printing of the authority file used by indexers, and the preparation of 3x5-inch catalog cards. This report discusses indexing and literature-searching routines, and the evolution of the computer retrieval system. Programs written for the computer, and the procedures for using them, are presented in Appendixes A through F.			

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Literature retrieval Program for library automation						

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, &, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those

imposed by security classification, using standard statements such as:

- (1) "Qualified requesters may obtain copies of this report from DDC."
- (2) "Foreign announcement and dissemination of this report by DDC is not authorized."
- (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."
- (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."
- (5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.